Deep learning to automate the assessment of cultural ecosystem services from social media data

Ana Sofia Cardoso[1], Ana Sofia Vaz[2], Francesco Renna[1]
[1]Instituto de Telecomunicações, Faculdade de Ciências da Universidade do Porto, Portugal
[2]Andalusian Inter-University Institute for Earth System Research (IISTA-CEAMA), University of Granada & Research Centre in Biodiversity and Genetic Resources (CIBIO-InBIO), University of Porto

Motivation

Cultural ecosystem services (CES) constitute the non-material benefits that people can experience from nature, such as recreation and ecotourism, as well as those pertaining to spiritual, religious, aesthetic or heritage values, among others [1].
An approach that combines different data from social media with advanced analytics, besides spatial analysis, remains underexplored in the context of CES assessment. Thus, the investment in methods that can identify features of ecosystems and nature through the content analysis of shared photos (or text), can constitute an asset to support the evaluation of CES, particularly, related to aesthetics and recreation or ecotourism [2].
In this work we studied automated image classification techniques using deep learning approaches to address CES.

Problem: Automated classification of social media photographs that can be useful for CES evaluation and for providing innovative solutions to the scientific community.

Methodologies proposed in the past

- Manual classification: questionnaires and interviews and social surveys [3].
- Convolutional neural networks (CNNs): capable of learning to identify similarities between patterns of information, in a manner that closely resembles a biological brain [4].

Proposed image classification methodology

We performed a classification of the content of photographs from the protected area "Peneda-Gerês" (Northern Portugal), that were withdrawn from the Flickr and Wikiloc social media platforms, specifying a time window of 2003-2017. This classification was based on "Nature" and "Human" labels.
The proposed image classification methods were evaluated over the dataset using a 5-fold-cross validation method, following the literature and taking into account the computational resources and the running time. The considered performance metrics (accuracy, sensitivity, specificity, and F1-score) were computed as the mean of the performance metrics obtained over the 5 different folds.

CNN architectures and transfer learning

Since we are coping with a small dataset, in order to improve the generalization of the model and avoid the overfitting, transfer learning and data augmentation schemes were considered.
To achieve that, two different CNNs architectures were implemented, the VGG16 and the ResNet152, in conjunction with two transfer learning scenarios (Places365 and ImageNet weights) and the weights obtained by training the networks from scratch, in order to verify the most appropriate and suitable for our study.
Regarding the details of the transfer learning strategy implemented, all the convolutional layers were kept frozen when training over our dataset, while the remaining 3 (for VGG16) and 1 (for ResNet152) fully connected layers were trained with our dataset.

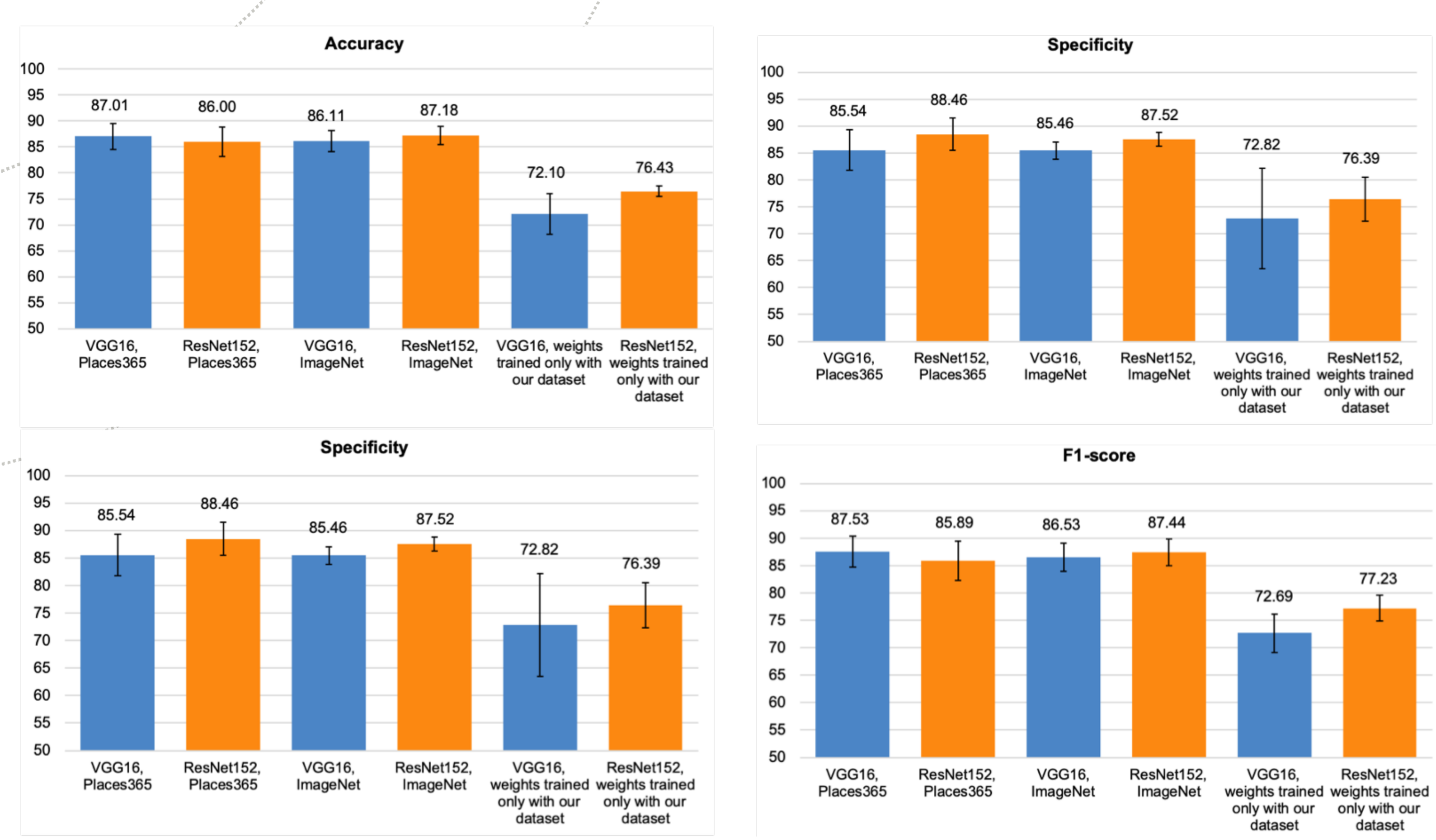
Data augmentation

Regarding data augmentation, 5 transformations (including horizontal flip, width shift, height shift and zoom) were implemented individually for each of the images in the training set. The images in the validation set were not included in this process, in order to avoid biased results.

Results

Nature vs. Human classification

Parameters:
- Learning rate: 0.0001 for ResNet152 and 0.000001 for VGG16.
- Mini batch size: 10
- Epochs: 50
- Optimizer: Adam



- VGG16 and ResNet152 with Places365 weights: p-value = 0.185
- VGG16 and ResNet152 with ImageNet weights: p-value = 0.303
- VGG16 and ResNet152 with the weights obtained by training the networks from scratch: p-value = 0.083
- VGG16 with Places365 and ImageNet weights: p-value = 0.516
- ResNet152 with Places365 and ImageNet weights: p-value = 0.287

Conclusions

- Overall, ResNet152 had a slightly finer performance than VGG16.
- ImageNet was the database where the two transfer learning scenarios achieved better results.
- The results showed that deep learning methods can offer significant contributions to assist in CES evaluation.
- Future work will focus on the improvement of the robustness of these models against scarcely labelled data via the use of semi-supervised approaches by leveraging autoencoder architectures and generative adversarial networks.

References and acknowledgements

[1] Assessment Millennium Ecosystem et al. *Ecosystems and human well-being* (Vol. 5). United States of America: Island press, 2005.
[2] Richards Daniel R.; Tunçer, Bige. Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosystem services*, 2018, 31: 318- 325. doi: 10.1016/j.ecoser.2017.09.004.
[3] Cheng Xin et al. Evaluation of cultural ecosystem services: A review of methods. *Ecosystem services*, 2019, *37:* 100925. doi: 10.1016/j.ecoser.2019.100925.
[4] Wang Jingya et al. Tracking natural events through social media and computer vision. In *Proceedings of the 24th ACM international conference on Multimedia.* 2016. p. 1097-1101. doi: 10.1145/2964284.2984067.

This work is funded by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020.